Forecasting of US GDP using Vector Auto Regressive Model

VEE Time Series Course (Students Project)

Fall 2011 Session

Nov 14th – Jan 8th 2012

Name: **Saurabh Khurana**

Email: pixelsallover.s@gmail.com

INDEX

INTRODUCTION

One of the most important ways to look at the overall economic aspect of any economy is through its Gross Domestic Product (GDP), which is the statistic used to measure the economy. We have considered the case for the U.S economy as the economy of the United States is the world's largest nominal economy. The U.S. economy maintains a very high level of output per person. Historically, the U.S. economy has maintained a stable overall GDP growth rate, a low unemployment rate, and high levels of research and capital investment funded by both national and, because of decreasing saving rates, increasingly by foreign investors. Since the 1960's, the United States economy absorbed savings from the rest of the world. The phenomenon is subject to discussion among economists.

A central feature of the U.S. economy is the economic freedom afforded to the private sector by allowing the private sector to make the majority of economic decisions in determining the direction and scale of what the U.S. economy produces. This is enhanced by relatively low levels of regulation and government involvement, as well as a court system that generally protects property rights and enforces contracts. From its emergence as an independent nation, the United States has encouraged science and invention.

The gross domestic product (GDP) or gross domestic income (GDI) is a measure of a country's overall economic output. It is the market value of all final goods and services made within the borders of a country in a year. It is often positively correlated with the standard of living though its use as a stand-in for measuring the standard of living has come under increasing criticism and many countries are actively exploring alternative measures to GDP for that purpose.

“Gross" means that GDP measures production regardless of the various uses to which that production can be put. Production can be used for immediate consumption, for investment in new fixed assets or inventories, or for replacing depreciated fixed assets.

GDP can be determined in three ways, all of which should in principle give the same result. They are the product (or output) approach, the income approach, and the expenditure approach. The most direct of the three is the product approach, which sums the outputs of every class of enterprise to arrive at the total.

The major advantage of GDP per capita as an indicator of standard of living is that it is measured frequently, widely and consistently. It is measured frequently in that most countries provide information on GDP on a quarterly basis, which allows a user to spot trends regularly.

The major disadvantage is that it is not, strictly speaking, a measure of standard of living. GDP is intended to be a measure of particular types of economic activity within a particular country. Nothing about the definition of GDP suggests it is necessarily a measure of standard of living. For instance, in an extreme example, a country which exported 100 per cent of its production and imported nothing would still have a high GDP, but a very poor standard of living. This concept is applicable to the prospective of the Indian Economy where although India has a high GDP but still there is no equitable distribution of gains among the population and hence it is still termed as a developing country.

To make sure that GDP can be most accurately compared year-to-year, the Bureau of Economic Analysis (BEA) usually reports *real GDP*.

How GDP Is Calculated:

To calculate real GDP, the BEA makes three important distinctions:

1. Imports and income from U.S. companies and people from outside the country are not included, so the impact of exchange rates and trade policies don't muddy up the number.
2. The effects of inflation are taken out.
3. Only the final product is counted, so that if someone in the U.S. makes shoelaces, and it is used to make shoes in the U.S. (there are a few companies left!) only the value of the shoe gets counted.

GDP is measured by the BEA quarterly. The BEA revises estimates as it receives better data throughout the next quarter

How GDP Affects the US Economy:

GDP is important for three reasons:

1. Most importantly, it is used to determine if the U.S. economy is growing more quickly or more slowly than the quarter before, or the same quarter the year before.
2. It is also used to compare the size of economies throughout the world.
3. It is to compare the relative growth rate of economies throughout the world.

Investors look at GDP growth to see if the economy is changing rapidly so they can adjust their asset allocation. In addition, investors compare country GDP growth rates to decide where the best opportunities are. Most investors like to purchase shares of companies that are in rapidly growing companies.

The Federal Reserve (Fed) uses the GDP growth rate as one of the indications of whether the economy needs to be restrained or stimulated

How GDP Affects You:

For example, if the GDP growth rate is speeding up, the Fed may raise interest rates to stem inflation. In this case, you would want to lock in a fixed-rate mortgage, because you know that an adjustable-rate mortgage will start charging higher rates next year.

If GDP is slowing down, or is negative, then you should dust off your resume. Declining GDP usually leads to layoffs and unemployment, but it can take several months. Declining GDP means business revenues are down. It can take awhile before executives can put together a layoff list and package. If you follow GDP statistics, you can be better prepared.

You could also use the GDP report from the BEA to look at which sectors of the economy are growing and which are declining. This would help you determine whether you should invest in, say, a tech-specific mutual fund vs a fund that focuses on agribusiness. It can also help you find training in sectors that are growing. Even during The Great Recession, healthcare related industries continued to hire.

Recent GDP Trends:

The 2000 recession was over by 2003, growing 2.5%, and expanding 3.6% in 2004. In 2005, Hurricane Katrina slowed growth to 2.9% in 2005. By March 2006, the economy recovered to 4.8% with the housing bubble peak. But high oil prices dragged on growth the rest of 2006, which only grew 2.7%. (Source: BEA, GDP News)

The economy only grew 1.2% in the first quarter of 2007 as the housing bubble popped. A falling dollar boosted exports, spurring growth to 3.2% in the second and 3.6% in the third quarters. When the Subprime Mortgage Crisis hit in October, growth slowed to 2.1%. Overall, the economy expanded 2.1% in 2007.

In 2008 and 2009, the economy contracted for four consecutive quarters. The last time this happened was during the Great Depression. The economy fell .7% in Q1 with the Stearns bailout, but resumed 1.5% growth by Q2. When the banking system imploded in the third quarter, the economy shrank 2.7%. The Lehman Brothers collapse delivered the death blow - the economy dropped 5.4% in Q4, growing only .4% for the year. GDP plummeted 6.4% in Q1 2009. By the second quarter, the economic stimulus package started to work: the economy shrank only .7% in Q2. It finally grew again by 2.2% in Q3.

The Bureau of Economic Analysis (BEA) promotes a better understanding of the U.S. economy by providing the most timely, relevant, and accurate economic accounts data in an objective and cost-effective manner. BEA produces economic accounts statistics that enable government and business decision-makers, researchers, and the American public to follow and understand the performance of the Nation's economy. To do this, BEA collects source data, conducts research and analysis, develops and implements estimation methodologies, and disseminates statistics to the public.

The Federal reserve bank of St. Loius provides statistical data for various research activities for the US economy. It aims at providing accurate and useful information on the current economic environment and the behavior of economic data around business cycle turning points.

The various market economic factors that tend to influence the market, hence the overall GDP growth rate for the US economy are

Long and Short Term Treasuries:

The U.S. Treasury issues many types of debt, from short-term to long-term. We can buy Treasury bonds that mature in three months or in thirty years, and quite a few in between. We can buy them straight from the Treasury or from others who are selling ones they bought from the Treasury. The interest rate on the bonds is the "yield to maturity" or just "yield," which accounts for both the coupon payments from the bonds and the price we paid for the bond. The yield curve simply plots the yield on the bond against its time to maturity.

Usually, the yield curve slopes up: longer-term bonds have higher yields than do short-term bonds, as people feel those longer-term bonds have more risk, requiring a higher return.

One of the most reliable and most often watched spreads is the one between 10-year U.S. Treasury bonds and 3-month T-bills. The natural and probably the most popular choice for growth is real GDP growth, taken at a quarterly or on a year-over-year basis, and predicting out one year.

An inverted yield curve has been viewed as an indicator of a pending economic recession. When short-term interest rates exceed long-term rates, market sentiment suggests that the long-term outlook is poor and that the yields offered by long-term fixed income will continue to fall. More recently, this viewpoint has been called into question as foreign purchases of securities issued by the U.S.

Current monetary policy has a significant influence on the yield curve spread and hence on real activity over the next several quarters. A rise in the short rate tends to flatten the yield curve as well as to slow real growth in the near term. This relationship, however, is only one part of the explanation for the yield curve's usefulness as a forecasting tool.

The yield curve for U.S. Treasury debt certificates is the one that investors use to predict the economy. Investors assume that the Treasury is the safest lender -- the least likely to default -- and therefore the rates on Treasury debt are least affected by risk.

This is one of the most important predictor variables affecting the GDP growth.

M1 Money Supply

The entire quantity of bills, coins, loans, credit and other liquid instruments in a country's economy is called money supply. The money supply is important to economists trying to understand how policies will affect interest rates and growth..A category of the money supply that includes all physical money such as coins and currency is denoted as M1 money Supply. It also includes demand deposits, which are checking accounts, and Negotiable Order of Withdrawal (NOW) Accounts. This is used as a measurement for economists trying to quantify the amount of money in circulation. The M1 is a very liquid measure of the money supply, as it contains cash and assets that can quickly be converted to currency.

These indices are always affecting the GDP growth in the sense that the market sentiments, industries, capital flows, investments are greatly influenced by these factors which in a way affect the overall economic growth.

MODEL

A Multivariate Time Series approach is used to model the GDP (the dependent Variable) with other variables as stated earlier (the independent variables)

The vector autoregressive (VAR) model is one of the most successful, flexible, and easy to use models for the analysis of multivariate time series. It is a natural extension of the univariate autoregressive model to dynamic multivariate time series. The VAR model has proven to be especially useful for describing the dynamic behavior of economic and financial time series and for forecasting. It often provides superior forecasts to those from univariate time series models and elaborate theory-based simultaneous equations models. Forecasts from VAR models are quite flexible because they can be made conditional on the potential future paths of specified variables in the model.

Vector auto regression (VAR) is being used as an econometric model is used to capture the evolution and the interdependencies between multiple time series, generalizing the univariate AR models. All the variables in a VAR are treated symmetrically by including for each variable an equation explaining its evolution based on its own lags and the lags of all the other variables in the model

A VAR model describes the evolution of a set of *k* variables (called ***endogenous variables***) over the same sample period (t = 1, ..., T) as a linear function of only their past evolution. The variables are collected in a $k \times 1$ vector y_t, which has as the i^{th} element $y_{i,t}$ the time t observation of variable y_i. For example, if the i^{th} variable is GDP, then $y_{i,t}$ is the value of GDP at t.

A ***(reduced) p-th order VAR***, denoted ***VAR(p)***, is

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + e_t,$$

For a general example of a VAR(p) with *k* variables

A VAR(1) in two variables can be written in matrix form (more compact notation) as

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix},$$

Or , equivalently, as the following system of two equations

$$y_{1,t} = c_1 + A_{1,1} y_{1,t-1} + A_{1,2} y_{2,t-1} + e_{1,t}$$

$$y_{2,t} = c_2 + A_{2,1} y_{1,t-1} + A_{2,2} y_{2,t-1} + e_{2,t}.$$

The lag length for the VAR(p) model may be determined using model selection criteria. The general approach is to fit VAR(p) models with orders p = 0,...,pmax and choose the value of p which minimizes some model selection criteria. One of the most common information criteria is the Akaike (AIC). The AIC criterion asymptotically overestimates the order with positive probability.

DATA SOURCES

The data used is based on a quarterly time series ranging from Quarter 3, 1959 to Quarter 1, 2009 (203 observations)

1. 10 year Treasury Bond Constant Maturity Rate (Percent)
   Source: www.research.stlouisfed.org
2. 3 Month Treasury Bill Secondary Market Rate (Percent)
   Source: www.research.stlouisfed.org
3. Gross Domestic Product (in Billion dollars) – quarterly data
   Source: www.research.stlouisfed.org
4. M1 Money Supply – quarterly data
   Source: www.research.stlouisfed.org

Data Frequency: Quarterly data being used in the analysis. If data is not quarterly available then monthly data is averaged out for the quarters as an estimate:

The log of GDP and M1 supply data are used in the entire analysis as to avoid any fluctuations and hence minimize the variance.

Although the independent variables could be taken in large quantities (such as M2, inflation rate) but a complex model would also give a model which a simple model would provide also.

The Entire Analysis is done using **SAS 9.1** and executing codes for the VAR procedure.

APPROACH & METHODOLOGY

STEP 1: Preparing and Plotting Time Series Data

Code:

```
title 'Analysis of U.S. Economic Variables';
   data Saurabh.D3;
     date=intnx( 'qtr', '01jan59'd, _n_-1 );
      format date yyq. ;
      input y1 y2 y3 y4;
      y1=log(y1);
      y2=log(y2);
      label y1='log(GDP)'
            y2='log(M1)'
                  y3='3 Month T-Bill Rate'
                  y4='10 Year T-Bill Rate';
   datalines;
508.5 139.1333    2.8167      4.0233
509.3 139.4667    3.0833      4.3500
```

………………………

; run;

```
proc timeseries data=Saurabh.D3 vectorplot=series;
      id date interval=qtr;
      var y1 y2 y3 y4;
ods graphics on;
run;
```

Output:

Simple Summary Statistics							
Variable	Type	N	Mean	Standard Deviation	Min	Max	Label
y1	Dependent	203	8.07510	1.06762	6.23147	9.58512	log(GDP)
y2	Dependent	203	6.20906	0.81267	4.92822	7.40845	log(M1)
y3	Dependent	203	5.31200	2.80339	0.11670	15.32670	3 Month T-Bill Rate
y4	Dependent	203	6.74499	2.57563	2.82330	15.13670	10 Year T-Bill Rate

Table 1.1

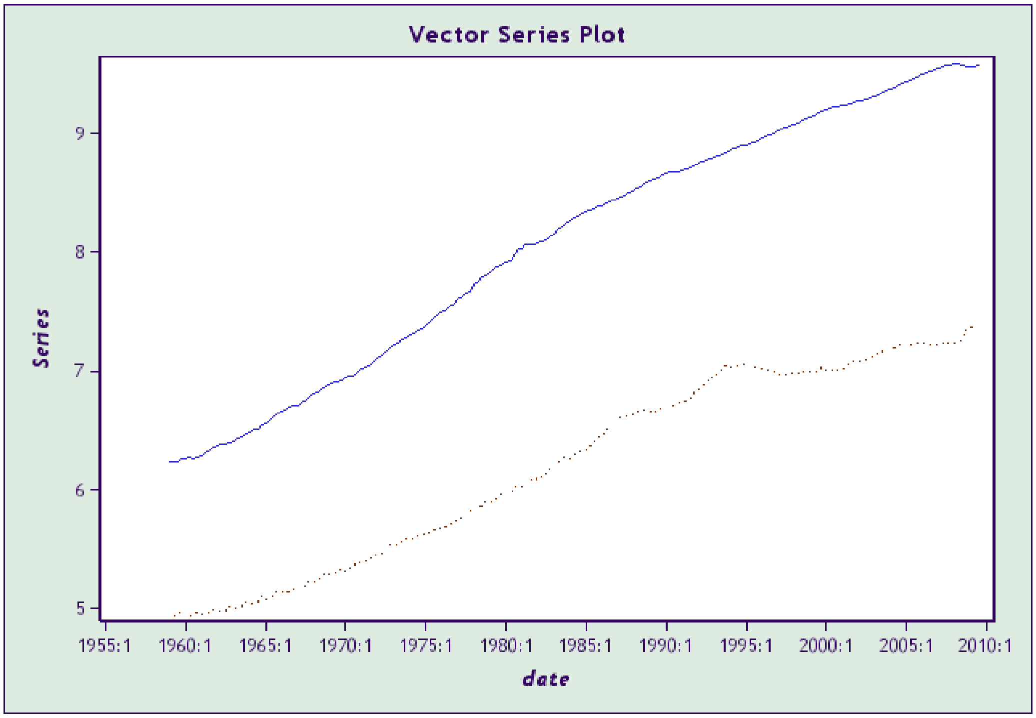


Figure 1.1

Shows the plot of the variables y1 and y2.

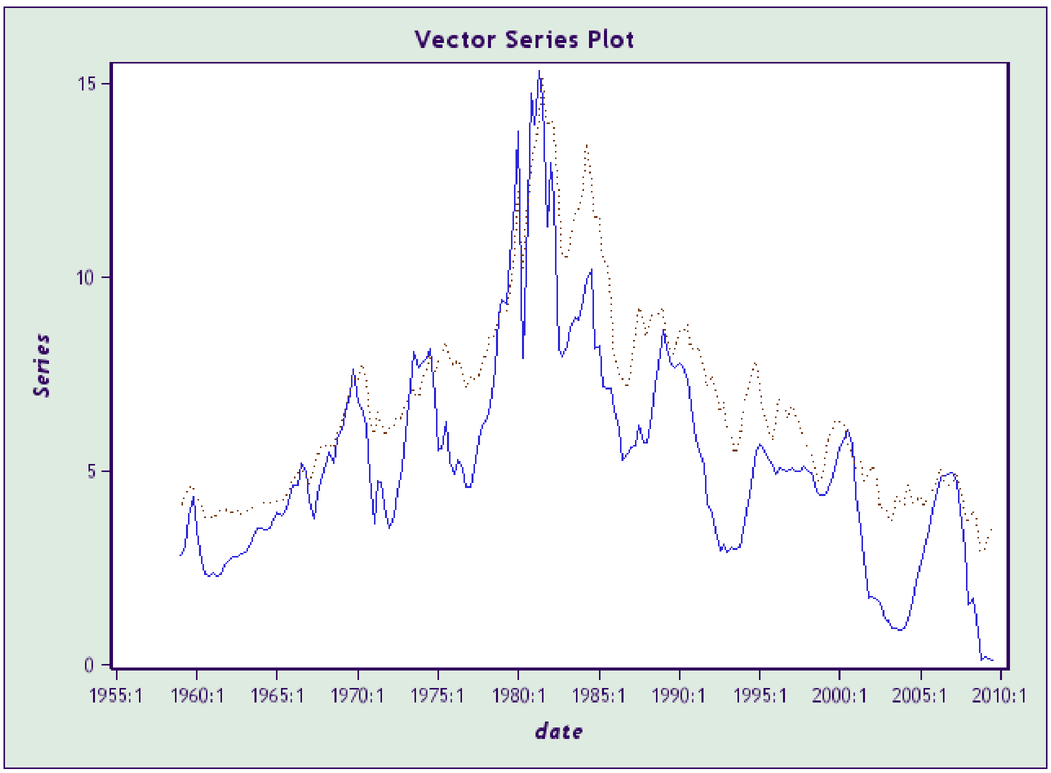


Figure 1.2

Shows the plot of the variables y3 and y4

Step 2: CHECK FOR STATIONARITY OF THE TIME SERIES

Dickey-Fuller Unit Root Test

Code:

```
proc varmax data=Saurabh.D3;
   model y1 y2 y3 y4/ p=1 dftest print=(roots);
run;
```

Output:

Dickey-Fuller Unit Root Tests					
Variable	Type	Rho	Pr < Rho	Tau	Pr < Tau
y1	Zero Mean	0.40	0.7800	6.75	0.9999
	Single Mean	-0.57	0.9213	-3.01	0.0360
	Trend	2.00	0.9997	1.37	0.9999
y2	Zero Mean	0.39	0.7771	7.45	0.9999
	Single Mean	-0.22	0.9450	-0.65	0.8557
	Trend	-1.36	0.9828	-0.67	0.9736
y3	Zero Mean	-2.45	0.2814	-1.13	0.2334
	Single Mean	-9.92	0.1329	-2.05	0.2646
	Trend	-10.93	0.3666	-2.25	0.4615
y4	Zero Mean	-0.75	0.5187	-0.64	0.4396
	Single Mean	-5.08	0.4243	-1.52	0.5230
	Trend	-5.20	0.8030	-1.57	0.8009

Table 2.1

The Table Shows the output for Dickey-Fuller tests for the nonstationarity of each series. The null hypothesis is to test a unit root. All series have a unit root.

Step 3: Finding Suitable Lag by finding minimum AIC

Code:

```
proc varmax data=Saurabh.D3;
model y1 y2 y3 y4 / minic=(type=aic p=5);
run;
```

Output:

Minimum Information Criterion						
Lag	**MA 0**	**MA 1**	**MA 2**	**MA 3**	**MA 4**	**MA 5**
AR 0	-2.980769	-2.86519	-2.928432	-3.056986	-3.12493	-3.179085
AR 1	-20.06813	-19.90738	-19.91975	-19.9188	-19.93381	-19.96348
AR 2	-20.24534	-20.02056	-20.01751	-20.01431	-20.08536	-20.11791
AR 3	-20.29117	-20.08845	-20.08075	-20.07453	-20.13407	-20.15642
AR 4	-20.43652	-20.25331	-20.23717	-20.19787	-20.3102	-20.49961
AR 5	-21.23953	-21.019	-21.06211	-20.98755	-21.12463	-21.14692

Table 3.1

The Table Shows Minimum AIC for the various VAR models of different lags from (p=0 to 5). The minimum AIC results for the VAR(5) model although there is not much significant difference from VAR(1).

STEP 4: Finding suitable Lag for the VAR model to be applied by testing Partial Auto Regression, Cross Correlation and Cannonical Correlations

Code:

```
proc varmax data=Saurabh.D3;
id date interval=qtr;
model y1 y2 y3 y4 / p=1 print=(parcoef pcorr pcancorr) lagmax=6 ;
run ;
```

Output:

Schematic Representation of Partial Autoregression						
Variable/Lag	***1***	***2***	***3***	***4***	***5***	***6***
y1	+...	....	....	....	-+..	....
y2	.+..	....	....	....	....	....
y3	..+.	....	....	..-.	....	..-.
y4	...+	....	....	....	....	....
+ is > 2*std error, - is < -2*std error, . is between						

Table 4.1

The table shows that the model can be obtained by an AR order *p*=1 since partial auto regression matrices are insignificant after lag 1 with respect to two standard errors.

Schematic Representation of Partial Cross Correlations						
Variable/Lag	***1***	***2***	***3***	***4***	***5***	***6***
y1	+...	....	....	....	-...	....
y2	.+..	....	...-	....	+-..	....
y3	..+.	....	....	..-.	....	..-.
y4	...+	....	....	....	....	...+
+ is > 2*std error, - is < -2*std error, . is between						

Table 4.2

The partial cross-correlation matrices in Figure 4.2 are insignificant after lag 1 with respect to two standard errors. This indicates that an AR order of *p*=1 can be an appropriate choice.

Partial Canonical Correlations							
Lag	Correlation1	Correlation2	Correlation3	Correlation4	DF	Chi-Square	Pr > ChiSq
1	0.98861	0.97466	0.97150	0.82479	16	717.38	<.0001
2	0.14014	0.11759	0.04109	0.00269	16	7.07	0.9719
3	0.29930	0.16243	0.10483	0.01890	16	25.46	0.0621
4	0.28151	0.22646	0.02368	0.00264	16	26.09	0.0528
5	0.52530	0.13323	0.06142	0.02124	16	58.99	<.0001
6	0.33617	0.18614	0.12083	0.01708	16	32.02	0.0099

Table 4.3

The Table shows that the partial canonical correlations between *Y(t)* and *Y(t-p).* After lag *p*=1, the partial canonical correlations are insignificant with respect to the 0.05 significance level, indicating than an AR order of *p*=1 would be an appropriate choice.

Step 5: TEST FOR EXOGENITY - Causality Test

Code:

```
proc varmax data=Saurabh.D3;
id date interval=qtr;
model y1 y2 y3 y4 / p=1;
causal group1=(y1) group2=(y2 y3 y4);
run;
```

Output:

A. Y1(**log(GDP))**

Granger-Causality Wald Test			
Test	DF	Chi-Square	Pr > ChiSq
1	3	22.38	<.0001

Test 1: Group 1 Variables:	y1
Group 2 Variables:	y2 y3 y4

Table 5.1

B. Y2(log(M1)

Granger-Causality Wald Test			
Test	DF	Chi-Square	Pr > ChiSq
1	3	34.44	<.0001

Test 1: Group 1 Variables:	y2
Group 2 Variables:	y1 y3 y4

Table 5.2

C. Y3(3 Month T-Bill Rate)

Granger-Causality Wald Test			
Test	DF	Chi-Square	Pr > ChiSq
1	3	2.27	0.5179

Test 1: Group 1 Variables:	y3
Group 2 Variables:	y1 y2 y4

Table 5.3

D. Y4(10 year T-Bill Rate)

Granger-Causality Wald Test			
Test	**DF**	**Chi-Square**	**Pr > ChiSq**
1	3	6.24	0.1004

Test 1: Group 1 Variables:	y4
Group 2 Variables:	y1 y2 y3

Table 5.4

The null hypothesis of the Granger-Causality test is that GROUP1 is influenced by itself, and not by GROUP2. If the test of hypothesis fails to reject the null, the variables in the GROUP1 may be considered as independent variables.

The Granger-Causality test statistics shows that Y1 (log(GDP)) is influenced by *Y2 , Y3, and Y4* . Hence, *Y1* should be treated as an endogenous variable, and the VAR(1) model is appropriate.

Step 6: Weak Exogeneity Test

Code:

```
proc varmax data=Saurabh.D3 outstat=bbb;
model y1-y4 / p=1 lagmax=2
ecm=(rank=1 normalize=y1);
cointeg rank=1 exogeneity;

run;
```

Output:

Testing Weak Exogeneity of Each Variables			
Variable	**DF**	**Chi-Square**	**Pr > ChiSq**
y1	1	20.62	<.0001
y2	1	23.79	<.0001
y3	1	0.48	0.4879
y4	1	1.65	0.1995

Table 6.1

The variable Y1 is not the weak exogeneity of other variables, Y2, Y3, and Y4

Therefore VAR(1) model is the appropriate choice. Hence VAR(1) model is applied to forecast the GDP.

Step 7: VAR(1) Model Fitting

Code:

```
proc varmax data=Saurabh.D3;
   model y1-y4 / p=1;
run;
```

Output:

Analysis of U.S. Economic Variables

The VARMAX Procedure

Type of Model	VAR(1)
Estimation Method	Least Squares Estimation

Constant Estimates	
Variable	***Constant***
y1	0.03792
y2	0.02665
y3	0.82392
y4	0.41008

Table 7.1

AR Coefficient Estimates					
Lag	***Variable***	***y1***	***y2***	***y3***	***y4***
1	***y1***	1.01229	-0.02038	-0.00146	0.00203
	y2	0.04222	0.94123	-0.00547	0.00573
	y3	0.02841	-0.15098	0.91036	0.05155
	y4	-0.01304	-0.01607	0.06485	0.91805

Table 7.2

Schematic Representation of Parameter Estimates		
Variable/Lag	C	AR1
y1	+	+--+
y2	+	++-+
y3	.	..+.
y4	.	...+
+ is > 2*std error, - is < -2*std error, . is between, * is N/A		

Table 7.3

Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
y1	CONST1	0.03792	0.00527	7.20	0.0001	1
	AR1_1_1	1.01229	0.00585	173.09	0.0001	y1(t-1)
	AR1_1_2	-0.02038	0.00775	-2.63	0.0092	y2(t-1)
	AR1_1_3	-0.00146	0.00057	-2.57	0.0108	y3(t-1)
	AR1_1_4	0.00203	0.00059	3.45	0.0007	y4(t-1)
y2	CONST2	0.02665	0.01044	2.55	0.0114	1
	AR1_2_1	0.04222	0.01158	3.65	0.0003	y1(t-1)
	AR1_2_2	0.94123	0.01535	61.33	0.0001	y2(t-1)
	AR1_2_3	-0.00547	0.00112	-4.88	0.0001	y3(t-1)
	AR1_2_4	0.00573	0.00116	4.92	0.0001	y4(t-1)
y3	CONST3	0.82392	0.51523	1.60	0.1114	1
	AR1_3_1	0.02841	0.57175	0.05	0.9604	y1(t-1)
	AR1_3_2	-0.15098	0.75775	-0.20	0.8423	y2(t-1)
	AR1_3_3	0.91036	0.05537	16.44	0.0001	y3(t-1)
	AR1_3_4	0.05155	0.05751	0.90	0.3712	y4(t-1)
y4	CONST4	0.41008	0.30548	1.34	0.1810	1
	AR1_4_1	-0.01304	0.33899	-0.04	0.9693	y1(t-1)
	AR1_4_2	-0.01607	0.44927	-0.04	0.9715	y2(t-1)
	AR1_4_3	0.06485	0.03283	1.98	0.0496	y3(t-1)
	AR1_4_4	0.91805	0.03410	26.92	0.0001	y4(t-1)

Table 7.4

Information Criteria	
AICC	-19.8479
HQC	-19.7204
AIC	-19.8529
SBC	-19.5254
FPEC	2.388E-9

Table 7.5

Cross Covariances of Residuals					
Lag	Variable	y1	y2	y3	y4
0	y1	0.00008	0.00000	0.00243	0.00111
	y2	0.00000	0.00030	-0.00083	-0.00062
	y3	0.00243	-0.00083	0.73364	0.30641
	y4	0.00111	-0.00062	0.30641	0.25790
1	y1	0.00002	-0.00002	0.00276	0.00175
	y2	0.00001	-0.00003	0.00135	0.00112
	y3	0.00031	-0.00089	0.05868	0.02028
	y4	0.00004	-0.00094	0.02739	0.02612
2	y1	0.00002	-0.00002	0.00150	0.00050
	y2	0.00000	0.00002	0.00003	0.00160
	y3	0.00028	-0.00071	-0.08780	-0.03966
	y4	0.00005	-0.00060	-0.03363	-0.01562
3	y1	0.00000	-0.00001	0.00163	0.00058
	y2	-0.00001	-0.00006	0.00065	0.00145
	y3	0.00063	-0.00086	0.18377	0.05436
	y4	0.00010	-0.00115	0.08665	0.03075
4	y1	0.00001	-0.00001	0.00076	0.00047
	y2	0.00000	0.00019	-0.00016	-0.00022
	y3	-0.00070	0.00002	-0.00474	0.01318
	y4	-0.00042	-0.00136	0.00482	-0.01066

Cross Covariances of Residuals					
Lag	Variable	y1	y2	y3	y4
5	y1	-0.00000	-0.00002	0.00014	-0.00016
	y2	0.00001	-0.00007	0.00021	0.00036
	y3	-0.00003	0.00008	0.04366	-0.03541
	y4	-0.00020	0.00039	0.00538	-0.05257
6	y1	0.00000	-0.00001	0.00118	-0.00007
	y2	-0.00000	-0.00001	0.00022	0.00079
	y3	-0.00013	-0.00156	0.04965	0.02765
	y4	0.00017	0.00013	0.02004	-0.01160
7	y1	-0.00000	0.00000	-0.00013	-0.00025
	y2	-0.00001	-0.00007	0.00106	0.00072
	y3	-0.00064	0.00015	-0.19722	-0.09521
	y4	-0.00022	0.00031	-0.07343	-0.04660
8	y1	-0.00000	0.00001	-0.00029	-0.00038
	y2	0.00001	0.00018	-0.00035	-0.00072
	y3	-0.00046	0.00041	-0.03892	-0.04404
	y4	-0.00010	-0.00029	-0.01426	-0.00459
9	y1	0.00001	-0.00001	-0.00091	-0.00062
	y2	0.00001	-0.00009	0.00027	0.00020
	y3	-0.00039	0.00058	0.08628	0.03747
	y4	0.00001	0.00049	0.02864	0.01167
10	y1	0.00001	0.00001	0.00079	0.00050
	y2	0.00000	-0.00003	-0.00061	0.00008
	y3	-0.00007	-0.00070	-0.07764	-0.02582
	y4	0.00029	0.00031	-0.02456	-0.00850
11	y1	0.00001	0.00001	0.00072	0.00036
	y2	-0.00001	-0.00008	0.00077	0.00029
	y3	-0.00062	0.00050	-0.11308	0.00014
	y4	-0.00016	0.00024	-0.02648	0.01324
12	y1	-0.00001	0.00002	-0.00062	0.00011
	y2	0.00001	0.00016	-0.00032	-0.00113

Cross Covariances of Residuals					
Lag	*Variable*	*y1*	*y2*	*y3*	*y4*
	y3	-0.00055	-0.00003	-0.05195	0.00199
	y4	-0.00007	-0.00080	0.00229	0.01540

Table 7.6

Cross Correlations of Residuals					
Lag	*Variable*	*y1*	*y2*	*y3*	*y4*
0	*y1*	1.00000	0.00307	0.32368	0.24837
	y2	0.00307	1.00000	-0.05607	-0.06987
	y3	0.32368	-0.05607	1.00000	0.70443
	y4	0.24837	-0.06987	0.70443	1.00000
1	*y1*	0.30838	-0.15303	0.36742	0.39380
	y2	0.08825	-0.09553	0.09067	0.12755
	y3	0.04161	-0.05995	0.07998	0.04663
	y4	0.00970	-0.10653	0.06297	0.10129
2	*y1*	0.21966	-0.10234	0.20031	0.11253
	y2	0.00986	0.05915	0.00179	0.18146
	y3	0.03688	-0.04767	-0.11967	-0.09118
	y4	0.01182	-0.06780	-0.07731	-0.06058
3	*y1*	0.06348	-0.04710	0.21673	0.13145
	y2	-0.07248	-0.19338	0.04364	0.16488
	y3	0.08391	-0.05811	0.25049	0.12498
	y4	0.02351	-0.13102	0.19919	0.11922
4	*y1*	0.11750	-0.04167	0.10066	0.10473
	y2	0.01189	0.61747	-0.01051	-0.02468
	y3	-0.09371	0.00162	-0.00646	0.03030
	y4	-0.09408	-0.15449	0.01108	-0.04132
5	*y1*	-0.02648	-0.14668	0.01882	-0.03534
	y2	0.05191	-0.22574	0.01410	0.04135
	y3	-0.00446	0.00570	0.05952	-0.08142

	y4	-0.04515	0.04383	0.01238	-0.20384
6	*y1*	0.01155	-0.05644	0.15703	-0.01575
	y2	-0.00830	-0.04277	0.01454	0.08933
	y3	-0.01695	-0.10485	0.06768	0.06356
	y4	0.03884	0.01459	0.04607	-0.04497
7	*y1*	-0.00997	0.02831	-0.01704	-0.05720
	y2	-0.06710	-0.23159	0.07105	0.08199
	y3	-0.08564	0.00986	-0.26882	-0.21887
	y4	-0.04910	0.03503	-0.16882	-0.18067
8	*y1*	-0.04434	0.09583	-0.03863	-0.08585
	y2	0.07738	0.58650	-0.02335	-0.08223
	y3	-0.06187	0.02742	-0.05305	-0.10126
	y4	-0.02188	-0.03259	-0.03277	-0.01781
9	*y1*	0.06644	-0.09196	-0.12115	-0.13964
	y2	0.05791	-0.28542	0.01791	0.02240
	y3	-0.05161	0.03877	0.11761	0.08615
	y4	0.00125	0.05511	0.06584	0.04526
10	*y1*	0.09774	0.05458	0.10484	0.11159
	y2	0.01509	-0.09718	-0.04098	0.00919
	y3	-0.00899	-0.04696	-0.10583	-0.05936
	y4	0.06480	0.03527	-0.05647	-0.03296
11	*y1*	0.11334	0.07653	0.09650	0.08100
	y2	-0.06103	-0.27402	0.05211	0.03314
	y3	-0.08208	0.03331	-0.15413	0.00032
	y4	-0.03537	0.02669	-0.06087	0.05135
12	*y1*	-0.06644	0.10465	-0.08260	0.02360
	y2	0.03575	0.54495	-0.02177	-0.12823
	y3	-0.07328	-0.00190	-0.07081	0.00457
	y4	-0.01465	-0.09059	0.00528	0.05971

Table 7.7

Univariate Model ANOVA Diagnostics				
Variable	R-Square	Standard Deviation	F Value	Pr > F
y1	0.9999	0.00887	720384	<.0001
y2	0.9995	0.01757	106703	<.0001
y3	0.9063	0.86733	476.25	<.0001
y4	0.9609	0.51424	1210.55	<.0001

Table 7.8

Univariate Model AR Diagnostics								
	AR1		AR2		AR3		AR4	
Variable	F Value	Pr > F	F Value	Pr > F	F Value	Pr > F	F Value	Pr > F
y1	21.34	<.0001	12.21	<.0001	8.94	<.0001	7.69	<.0001
y2	1.83	0.1771	1.17	0.3125	3.16	0.0257	39.47	<.0001
y3	1.28	0.2590	2.26	0.1073	7.04	0.0002	5.66	0.0003
y4	2.06	0.1523	1.53	0.2189	2.25	0.0840	1.99	0.0969

Table 7.9

FORECAST VALUES FOR A 6 QUARTERS AHEAD

Forecasts						
Variable	Obs	Time	Forecast	Standard Error	95% Confidence Limits	
y1	204	2009:4	9.59020	0.00887	9.57281	9.60759
	205	2010:1	9.60109	0.01255	9.57649	9.62570
	206	2010:2	9.61149	0.01545	9.58120	9.64177
	207	2010:3	9.62146	0.01798	9.58622	9.65671
	208	2010:4	9.63109	0.02030	9.59130	9.67088
	209	2011:1	9.64043	0.02248	9.59637	9.68449
y2	204	2009:4	7.42329	0.01757	7.38886	7.45772
	205	2010:1	7.43633	0.02439	7.38851	7.48414
	206	2010:2	7.44785	0.02960	7.38984	7.50585
	207	2010:3	7.45810	0.03403	7.39141	7.52480
	208	2010:4	7.46731	0.03799	7.39286	7.54177
	209	2011:1	7.47565	0.04159	7.39414	7.55716
y3	204	2009:4	0.26513	0.86733	-1.43480	1.96507
	205	2010:1	0.38969	1.18587	-1.93457	2.71395
	206	2010:2	0.49670	1.40713	-2.26122	3.25462
	207	2010:3	0.58873	1.57717	-2.50248	3.67993
	208	2010:4	0.66797	1.71457	-2.69253	4.02846
	209	2011:1	0.73627	1.82910	-2.84870	4.32124
y4	204	2009:4	3.35031	0.51424	2.34241	4.35821
	205	2010:1	3.25862	0.72656	1.83459	4.68266
	206	2010:2	3.18218	0.88930	1.43919	4.92518
	207	2010:3	3.11862	1.02618	1.10734	5.12991
	208	2010:4	3.06595	1.14619	0.81946	5.31244
	209	2011:1	3.02245	1.25379	0.56506	5.47984

Table 7.10

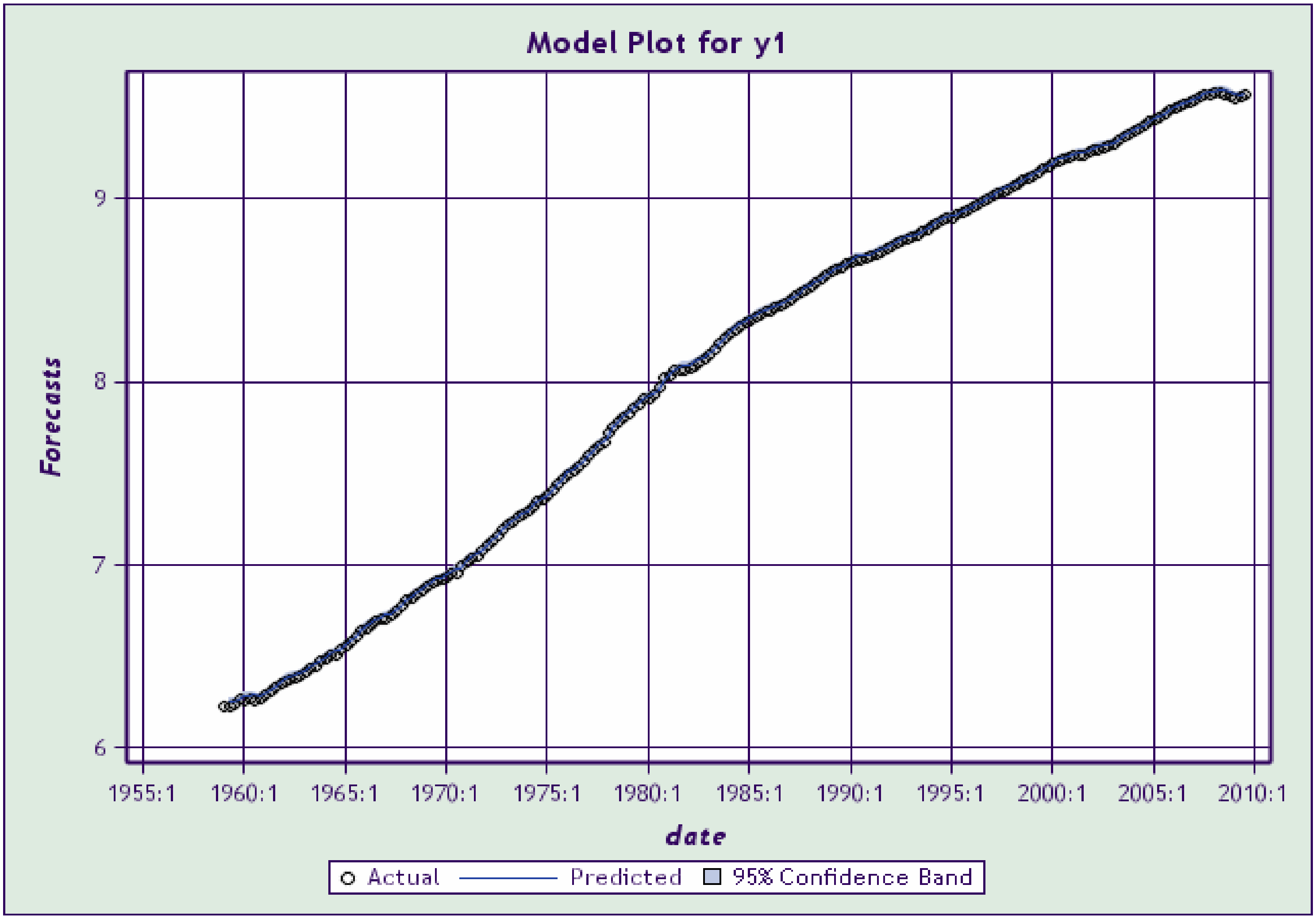

Model Plot for y1
Forecasts
9
8
7
6
1955:1
1960:1
1965:1
1970:1
1975:1
1980:1
1985:1
1990:1
1995:1
2000:1
2005:1
2010:1
date
Actual
Predicted
95% Confidence Band

Figure 7.1

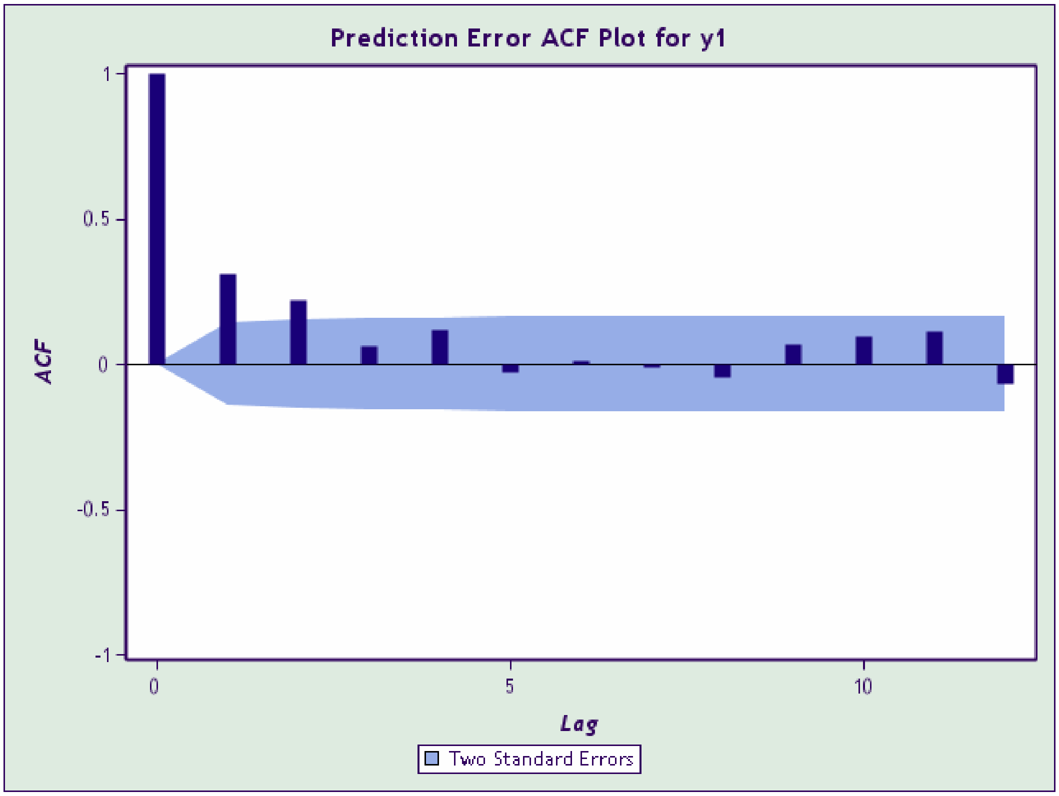


Figure 7.2

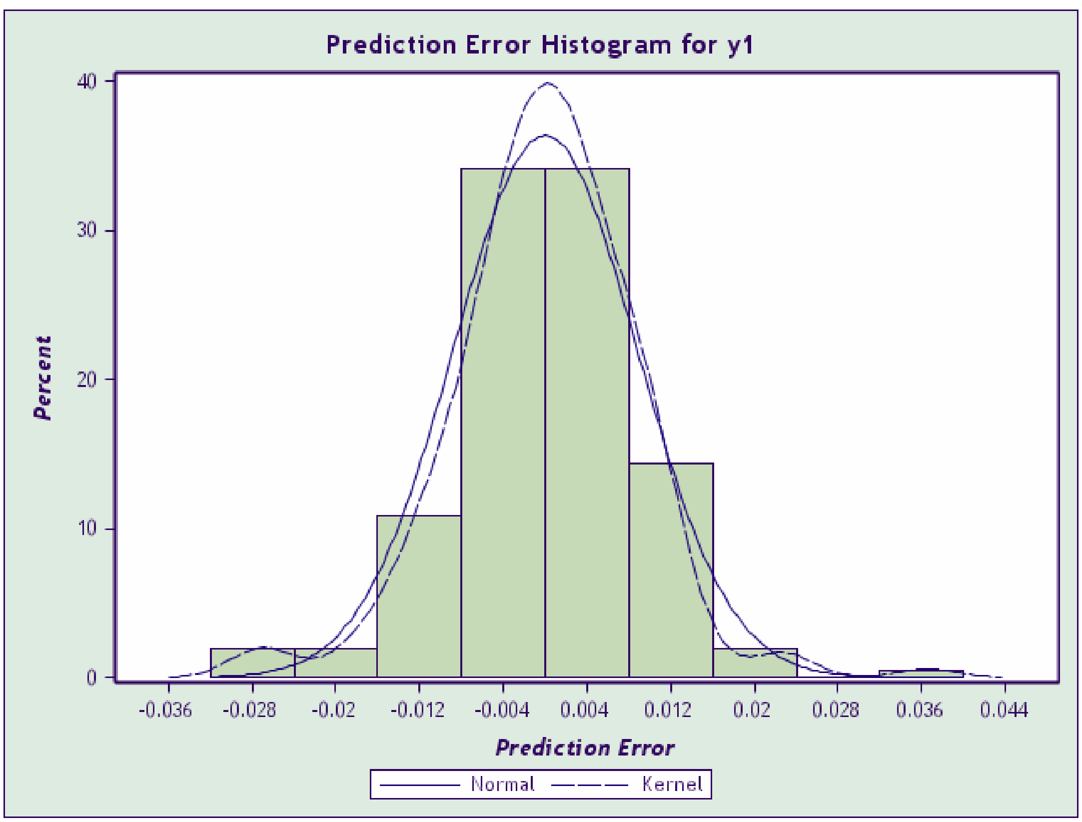


Figure 7.3

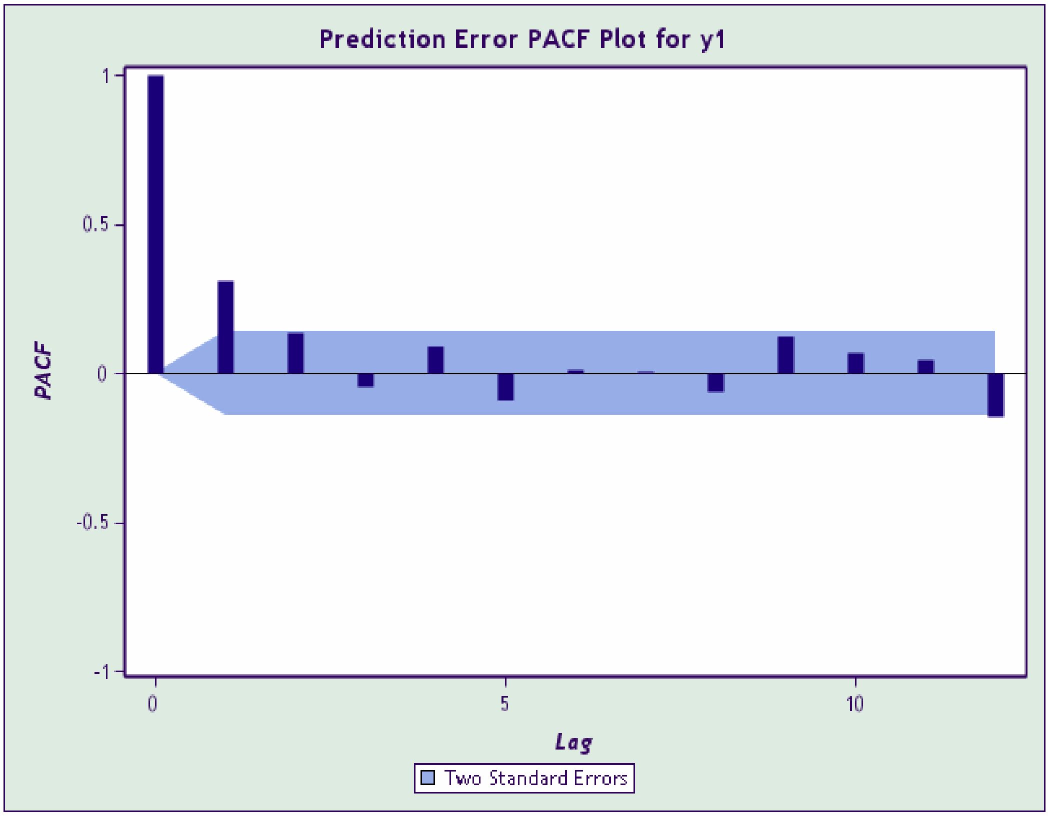


Figure 7.4

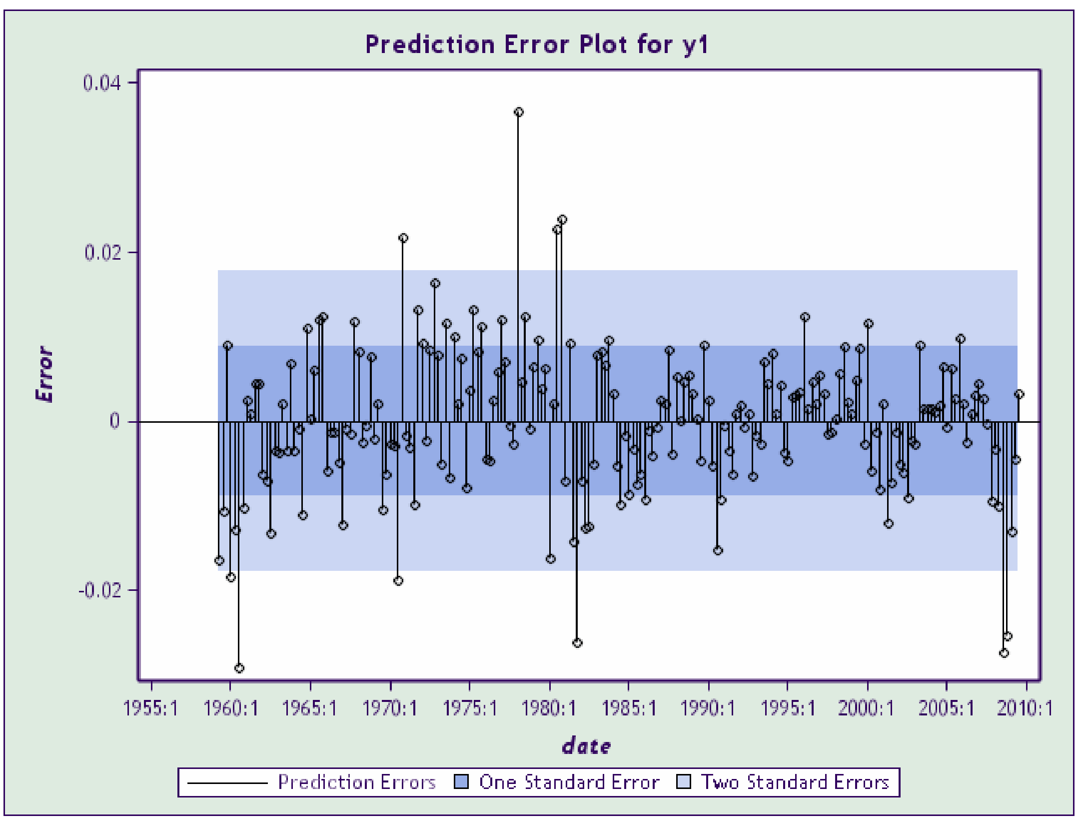


Figure 7.5

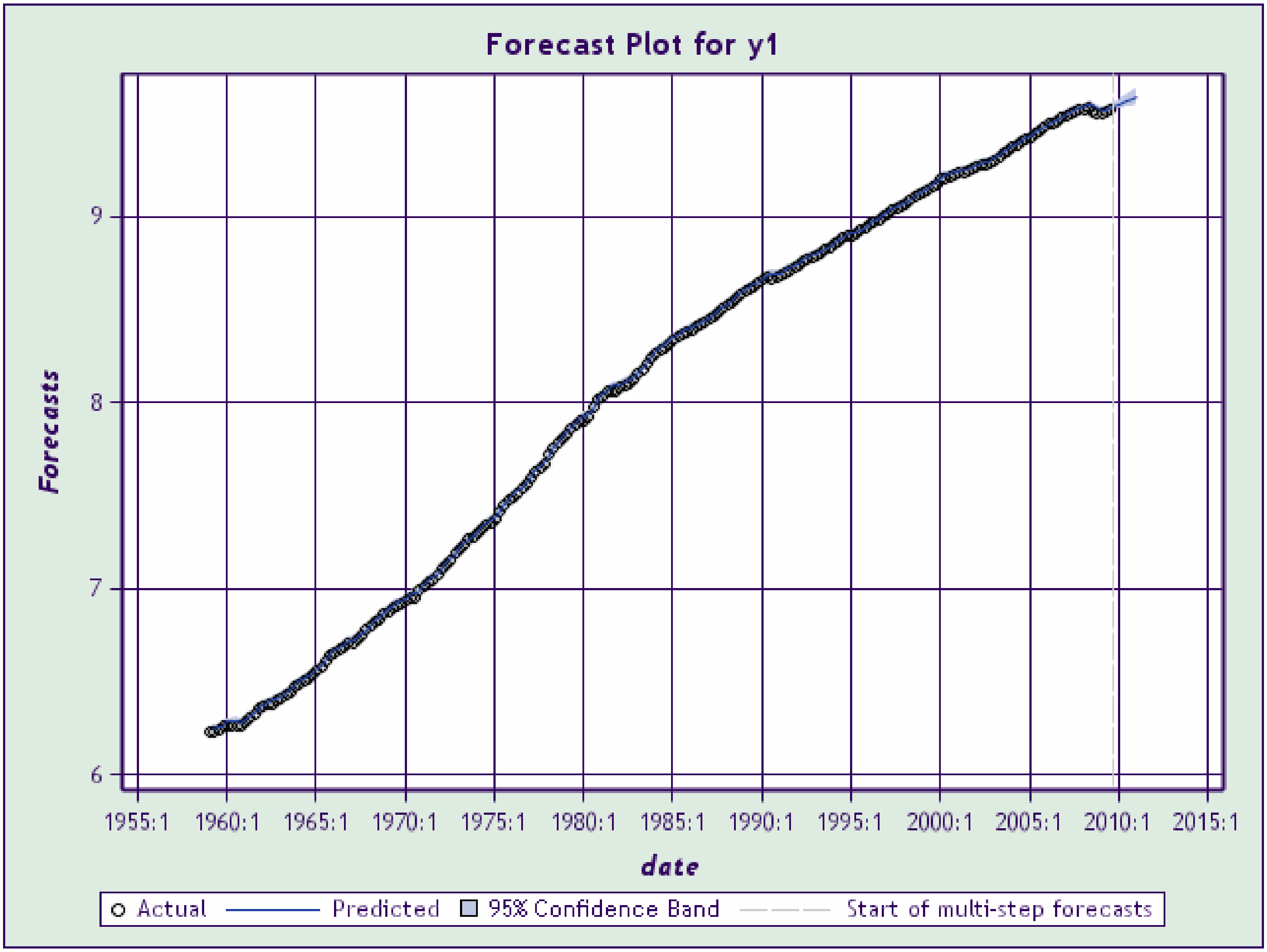
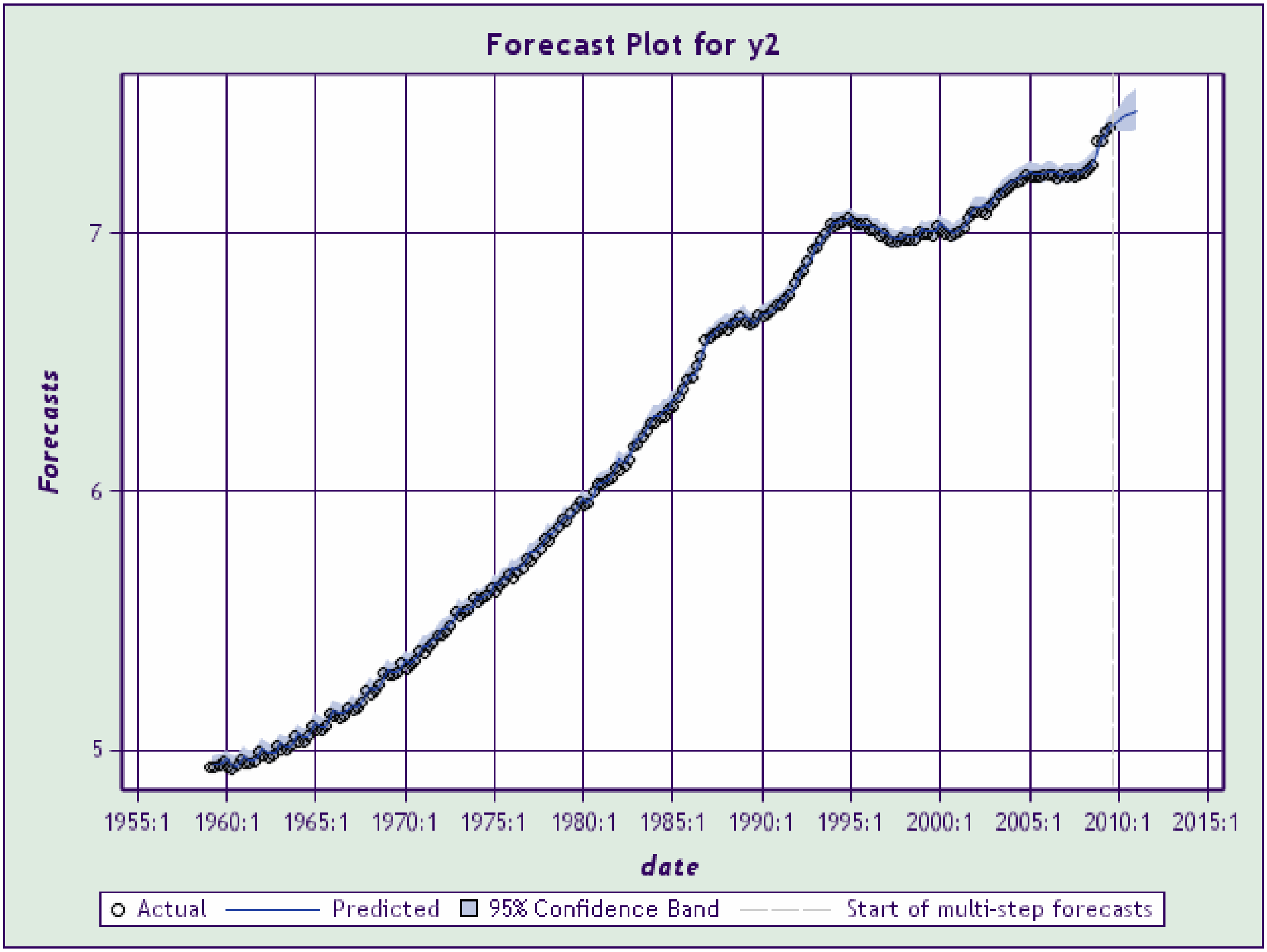

Forecast Plot for y1
Forecasts
9
8
7
6
1955:1
1960:1
1965:1
1970:1
1975:1
1980:1
1985:1
1990:1
1995:1
2000:1
2005:1
2010:1
2015:1
date
Actual
Predicted
95% Confidence Band
Start of multi-step forecasts

Figure 7.6



Forecast Plot for y2
Forecasts
7
6
5
1955:1
1960:1
1965:1
1970:1
1975:1
1980:1
1985:1
1990:1
1995:1
2000:1
2005:1
2010:1
2015:1
date
Actual
Predicted
95% Confidence Band
Start of multi-step forecasts

Figure 7.7

FORECAST VALUES FOR A 12 QUARTERS AHEAD

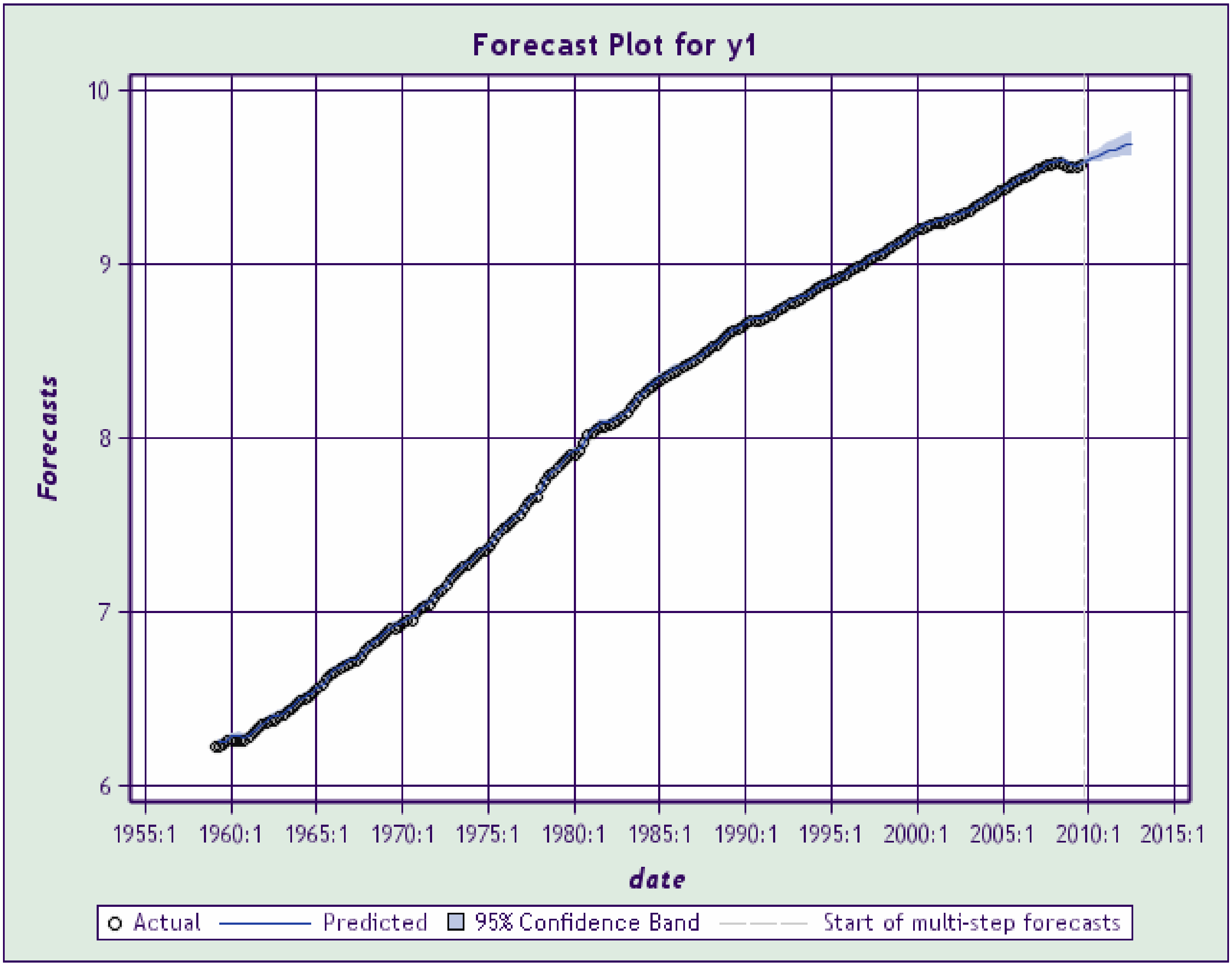


Figure 7.8

Forecasts						
Variable	**Obs**	**Time**	**Forecast**	**Standard Error**	**95% Confidence Limits**	
y1	204	2009:4	9.59020	0.00887	9.57281	9.60759
	205	2010:1	9.60109	0.01255	9.57649	9.62570
	206	2010:2	9.61149	0.01545	9.58120	9.64177
	207	2010:3	9.62146	0.01798	9.58622	9.65671
	208	2010:4	9.63109	0.02030	9.59130	9.67088
	209	2011:1	9.64043	0.02248	9.59637	9.68449
	210	2011:2	9.64952	0.02456	9.60137	9.69766
	211	2011:3	9.65841	0.02658	9.60631	9.71050
	212	2011:4	9.66713	0.02854	9.61118	9.72307
	213	2012:1	9.67571	0.03047	9.61599	9.73543
	214	2012:2	9.68418	0.03237	9.62074	9.74762
	215	2012:3	9.69256	0.03425	9.62544	9.75968

Table 8.11

RESULTS AND DISCUSSION

1. The Dickey Fuller Unit Root test shows that all the series are stationary and hence no differencing is required. . All series have a unit root. (Table 2.1)

2. Minimum AIC for various VAR models was found and from the table 3.1 we infer that the model AR(5) has the minimum AIC. Although there is not much significant difference from VAR(1). So further tests are required to select the appropriate lag model.

3. The table 4.1 shows that the model can be obtained by an AR order *p*=1 since partial auto regression matrices are insignificant after lag 1 with respect to two standard errors.

4. The partial cross-correlation matrices in Table 4.2 are insignificant after lag 1 with respect to two standard errors. This indicates that an AR order of *p*=1 can be an appropriate choice.

5. The Table 4.3 shows that the partial canonical correlations between *Y(t)* and *Y(t-p).* After lag *p*=1, the partial canonical correlations are insignificant with respect to the 0.05 significance level, indicating than an AR order of *p*=1 would be an appropriate choice.

6. Table 5.1 for the Granger-Causality test statistics shows that Y1 (log(GDP)) is influenced by Y2, Y3, and Y4 . Hence, Y1 should be treated as an endogenous variable, and the VAR(1) model is appropriate.

7. The variable Y1(log(GDP)) is not the weak exogeneity of other variables, Y2, Y3, and Y4 as shown in the Table 6.1

Through all the tests above we infer that the series is stationary and VAR(1) model can be applied. We apply the VAR(1) model.

Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
y1	CONST1	0.03792	0.00527	7.20	0.0001	1
	AR1_1_1	1.01229	0.00585	173.09	0.0001	y1(t-1)
	AR1_1_2	-0.02038	0.00775	-2.63	0.0092	y2(t-1)
	AR1_1_3	-0.00146	0.00057	-2.57	0.0108	y3(t-1)
	AR1_1_4	0.00203	0.00059	3.45	0.0007	y4(t-1)

AIC	-19.8529

Forecasts Values for 6 Quarters ahead						
Variable	Obs	Time	Forecast	Standard Error	95% Confidence Limits	
y1	204	2009:4	9.59020	0.00887	9.57281	9.60759
	205	2010:1	9.60109	0.01255	9.57649	9.62570
	206	2010:2	9.61149	0.01545	9.58120	9.64177
	207	2010:3	9.62146	0.01798	9.58622	9.65671
	208	2010:4	9.63109	0.02030	9.59130	9.67088
	209	2011:1	9.64043	0.02248	9.59637	9.68449

Forecasts Values for 12 quarters ahead						
Variable	Obs	Time	Forecast	Standard Error	95% Confidence Limits	
y1	204	2009:4	9.59020	0.00887	9.57281	9.60759
	205	2010:1	9.60109	0.01255	9.57649	9.62570
	206	2010:2	9.61149	0.01545	9.58120	9.64177
	207	2010:3	9.62146	0.01798	9.58622	9.65671
	208	2010:4	9.63109	0.02030	9.59130	9.67088
	209	2011:1	9.64043	0.02248	9.59637	9.68449
	210	2011:2	9.64952	0.02456	9.60137	9.69766
	211	2011:3	9.65841	0.02658	9.60631	9.71050
	212	2011:4	9.66713	0.02854	9.61118	9.72307
	213	2012:1	9.67571	0.03047	9.61599	9.73543

Forecasts Values for 12 quarters ahead						
Variable	Obs	Time	Forecast	Standard Error	95% Confidence Limits	
	214	2012:2	9.68418	0.03237	9.62074	9.74762
	215	2012:3	9.69256	0.03425	9.62544	9.75968

FUTURE SCOPE

The project right now takes into account only a few variables which are governed and regulated by the US government in order to govern the economy stably. There are many other factors that could also be taken into account such as inflation, unemployment rate and other indices (M2, CPI, CLI) as well in order to improvise the model and getting better estimates. Similar methodology could be developed to forecast other variables such as Inflation.

REFERENCES

1. Camba-Mendez, Gonzalo, George Kapetanios, Richard J. Smith, and Martin R. Weale. 'An Automatic Leading Indicator of Economic Activity: Forecasting GDP Growth for European Countries', Econometrics Journal 4, (2001), S56-S90.
2. Carpenter, Seth B. and Selva Demiralp (2008) "Money and the Transmission of Monetary Policy." unpublished manuscript, Board of Governors of the Federal Reserve System and Koç University.
3. Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin and G. Rünstler (2008a), Short term forecasts of euro area GDP. ECB Working Paper series, forthcoming.
4. Dickey, D.A. and W.A. Fuller. (1979) "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," Journal of the American Statistical Association 74(2), 427 - 431.
5. Ang, A. and M. Piazzesi, 2003, A No-Arbitrage Vector autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables, Journal of Monetary Economics 50, 4, 745-787.
6. www.stloiusfed.org, Federal Reserve Bank of St Louis
7. www.nber.org, National Bureau of Economic Research
8. www.ustreas.gov United States Department of Treasury
9. www.wikipedia.com
10. www.investopedia.com
11. www.bea.gov